Electromagnetism and Quantum Theory

Luis Fernando Alday*

Part1: Electromagnetism

This half of the course is a short introduction to the mathematical description of the salient points of the theory of electromagnetism. I will assume that you have some familiarity with the physical phenomena of electromagnetism which the theory aims to encompass and account for. For many of you this will have come from physics courses in school, and can be reinforced by consulting a book like Duffin (see below). Our main concern is with the formalism: The language of vector operators and the integral theorems of Gauss, Green and Stokes are precisely tailored to handle electromagnetic theory. Once one has the theory written that way, there are various tantalising hints of how to carry it forward into the four-dimensional space-time formalism of special relativity - but that is another course.

Synopsis

Maxwell's equations in vacuum with sources. Lorentz force law. Plane waves and polarization. Electrostatics and Magnetostatics. Energy density and the Poynting vector. Scalar and vector potentials. Gauge invariance. Maxwell's equations in the Lorentz gauge. The wave equation for potentials.

Reading list

One good source, with extensive discussion of the experimental background, which I shall of necessity by omitting is:

W.J. Duffin, Electricity and Magnetism, McGraw-Hill (2001), Fourth Edition, Chapters 1-4,7,8,13.

Classic texts on electromagnetism:

R. Feynman, Lectures in Physics, Vol.2. Electromagnetism, Addison Wesley.

J.D. Jackson, Classical Electrodynamics, John Wiley.

---

*Largely based on notes by Paul Tod.

1 Time independent Electromagnetic Theory

1.1 Electrostatics

This is the subject of static electricity, which we've all played with: recall rubbing a pen on your cat and picking up bits of paper. We interpret this as charging up or adding charge to the pen. As greeks didn't have pens, they used amber. The greek word for amber is electron and this is where the subject gets its name from.

We will assume the following facts about charge:

- there exist charges in nature;
- charge can be positive or negative;
- charge cannot be created or destroyed;

For a more sophisticated standpoint, we may know that charge is discrete, and that electrons have charge -1 and protons charge +1 in suitable units, but for our purposes we want to think of charge as continuous. As a mathematical idealization, we represent charge by a real function, the *charge density* $\rho(x, y, z, t)$, with the property that the total charge in a region V of space is

$$Q = \int_V \rho dV = \iiint_V \rho \; dxdydz. \tag{1}$$

Other idealizations are possible, and we shall sometimes use them. For example, it may be convenient to think of charges residing on a surface, something two-dimensional, and then *charge surface density* $\sigma(x, y, t)$ is charge per unit area, so that the charge on a piece S of surface is

$$Q = \int_S \sigma dS; \tag{2}$$

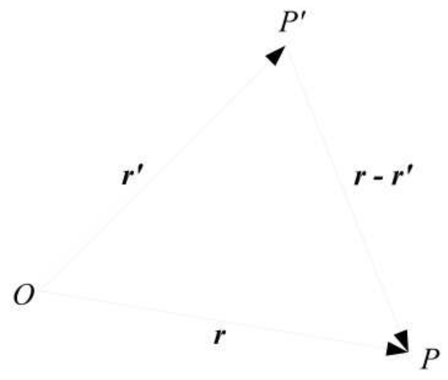
similarly, for a one-dimensional distribution of charge, *charge line density* is charge per unit length on some curve. Finally, a *point charge* corresponds to a finite, nonzero charge localized at a point.

How do charges interact with each other? An important experimental fact about point charges is the

Coulomb's Law: Given two point particles P, P' with charges Q, Q' at positions $\vec{r}, \vec{r}'$, the electric force on P due to P' is

$$\vec{F} = \frac{1}{4\pi\epsilon_0} QQ' \frac{\vec{r} - \vec{r}'}{|\vec{r} - \vec{r}'|^3} \tag{3}$$

where ϵ_0 is a constant and accounts for the correct units.



We treat this law as an experimental fact. To see how one might have discovered it, note that:

- it is an inverse-square law, acting along the line joining the particles;
- it is proportional to the product of the charges, so that like charges repel and unlike charges attract.

It is convenient to think of this law as follows: P' generates an *electric field* to which P responds.

Put P' at the origin and define the electric field at $\vec{r}$ to be

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} Q' \frac{\vec{r}}{r^3} \tag{4}$$

Then P, which has charge Q, when placed in this field is subject to (or "feels") the force

$$\vec{F} = Q\vec{E} \tag{5}$$

Replace P' by various particles $P_1, P_2, ..., P_n$, at points $\vec{r}_1, \vec{r}_2, ..., \vec{r}_n$ with charges $Q_1, Q_2, ..., Q_n$. The forces will add as vectors so if we defined the electric field now as

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{n} \frac{Q_i(\vec{r} - \vec{r}_i)}{|\vec{r} - \vec{r}_i|^3}, \tag{6}$$

then, as before, P placed at $\vec{r}$ is subject to the force

$$\vec{F}(\vec{r}) = Q\vec{E}(\vec{r}) \tag{7}$$

We want to explore (6). Note first that it is the gradient of another function:

$$\vec{E} = -\nabla\Phi, \quad \Phi = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{n} \frac{Q_i}{|\vec{r} - \vec{r}_i|}. \tag{8}$$

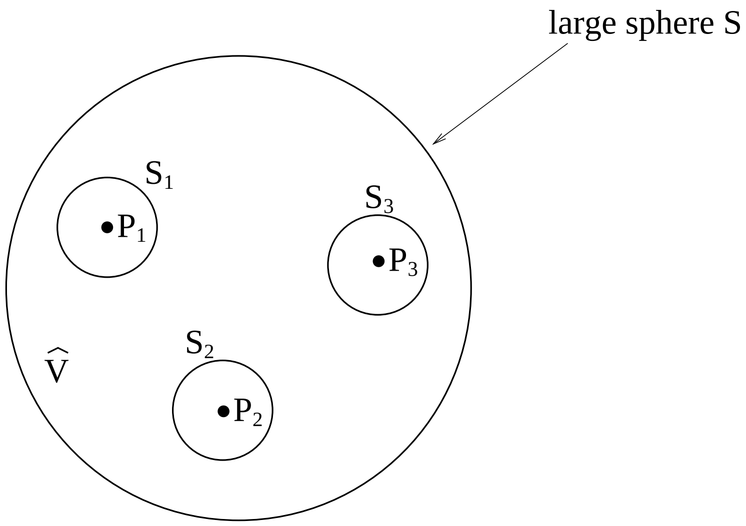
This function Φ is called the *electric potential*. In particular, therefore

$$\nabla \wedge \vec{E} = 0, \tag{9}$$

so the force (7) is a *conservative force* in the language of mechanics (familiar to you from Mods). We also have

$$\nabla \cdot \vec{E} = \nabla^2 \Phi = 0, \qquad \text{except at } \vec{r} = \vec{r}_i \tag{10}$$

Consider the diagram



i.e. S is a large sphere containing all P_i, S_i is a small sphere containing only P_i and $\hat{V}$ is the region in between.

Then $\nabla \cdot \vec{E} = 0$ on $\hat{V}$, so

$$\int_S \vec{E} \cdot d\mathbf{S} - \sum_{i=1}^{n} \int_{S_i} \vec{E} \cdot d\mathbf{S} = \int_{\hat{V}} \nabla \cdot \vec{E} dV = 0 \tag{11}$$

the first two are surface integrals (with $d\mathbf{S}$ pointing away from the center of the corresponding sphere), while the last one is a volume integral. Make sure you understand the signs in this equation. Let us focus in the integral over S_1, taking $\vec{r}_1 = 0$ for simplicity we obtain:

$$\begin{aligned} \int_{S_1} \vec{E} \cdot d\mathbf{S} &= \int_{S_1} \frac{1}{4\pi\epsilon_0} \left(Q_1 \frac{\vec{r}}{r^3} \cdot \mathbf{n} + \text{terms integrating to zero} \right) dS \\ &= \frac{1}{4\pi\epsilon_0} Q_1 \int \frac{1}{r^2} r^2 \sin\theta d\theta d\phi \\ &= \frac{1}{\epsilon_0} Q_1, \end{aligned}$$

in the second step we used $\mathbf{n} = \frac{\vec{r}}{r}$. Each term in the sum (11) can be calculated like this, so we obtain:

$$\int_S \vec{E} \cdot d\mathbf{S} = \sum_{i=1}^{n} \int_{S_i} \vec{E} \cdot d\mathbf{S} = \frac{1}{\epsilon_0}(Q_1 + Q_2 + ... + Q_n) = \frac{1}{\epsilon_0} \times \text{total charge inside S} \tag{12}$$

This is the *Gauss's Law* in its integral form:

Gauss's Law: The flux of $\vec{E}$ out of $V = \frac{1}{\epsilon_0} \times$ total charge in V.

If we have a smoothed out charge density instead of point charges, Gauss's law would read

$$\int_S \vec{E} \cdot d\mathbf{S} = \frac{1}{\epsilon_0} \int_V \rho dV, \tag{13}$$

where V is the volume inside the closed surface S. We will assume this to be true.

Now, by Divergence theorem we have

$$\int_V \left(\nabla \cdot \vec{E} - \frac{1}{\epsilon_0}\rho \right) dV = 0 \tag{14}$$

but if this is to hold for all possible regions V, then

$$\nabla \cdot \vec{E} = \frac{1}{\epsilon_0}\rho, \tag{15}$$

which is the *differential version of Gauss's Law.*

This together with $\vec{E} = -\nabla\Phi$ implies *Poisson's equation*

$$\nabla^2\Phi = -\frac{1}{\epsilon_0}\rho \tag{16}$$

What is the solution for this equation? remember that for point particles we had (see (8))

$$\Phi = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{n} \frac{Q_i}{|\vec{r} - \vec{r}_i|} \tag{17}$$

so we might guess that the solution of (16) is

$$\Phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \iiint_V \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} dx'dy'dz' \tag{18}$$

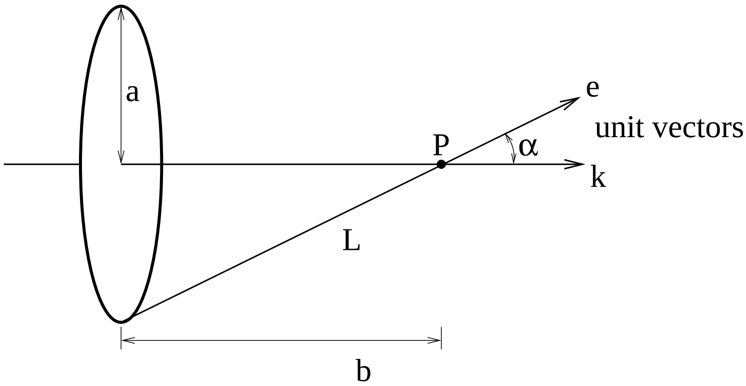
where V is the region in which $\rho \neq 0$. In fact, this is the correct guess.

An example

The main problem of Electrostatics is to obtain $\vec{E}$ given the charge distribution. We have solved that problem with (18): Given the density ρ, integrate this to find Φ and then

$\vec{E} = -\nabla\Phi$. However, in simple cases we can go straight to $\vec{E}$ by using (4). Here is an example:

Total charge Q is spread out uniformly round a plane circular wire of radius a. Find the electric field at a point P on the axis of the circle, at a distance b from the centre.



We can do this one directly from the inverse square law: $\vec{E} = \frac{q}{4\pi\epsilon_0}\frac{\vec{r}}{r^3}$.

Cut the circular wire into elements, each of legth $ad\theta$; then each contains charge equal to $\frac{Q}{2\pi}d\theta$ and so contributes $\frac{Q}{2\pi}d\theta\frac{1}{4\pi\epsilon_0}\frac{\vec{e}}{L^2}$, (with $\vec{e}$ as in the figure) to $\vec{E}$ at P. Adding these up around the circle leads, by symmetry, to a vector along $\vec{k}$. Note, from the diagram, that $\vec{e}\cdot\vec{k} = \cos\alpha$, so that $\vec{E}(p) = E\vec{k}$, with E a scalar and given by

$$\begin{aligned} E &= \int \frac{Q}{2\pi}\cdot d\theta \cdot \frac{1}{4\pi\epsilon_0}\cdot\frac{\cos\alpha}{L^2} \\ &= \frac{Q}{4\pi\epsilon_0}\frac{\cos\alpha}{L^2} \\ &= \frac{Q}{4\pi\epsilon_0}\frac{b}{(a^2+b^2)^{3/2}} \end{aligned}$$

which is the answer.

It is trickier to find $\vec{E}$ for a point P' not on the axis.

Φ and the Potential Energy

Let us think about Newton's equations of motion for a charged particle of mass m and charge Q subject to the force (7)

$$m\vec{a} = \vec{F} = Q\vec{E} = -Q\nabla\Phi \tag{19}$$

Then

$$\frac{d}{dt}\left(\frac{1}{2}m\vec{v}\cdot\vec{v}\right) = m\vec{v}\cdot\vec{a} = -Q\vec{v}\cdot\nabla\Phi = -Q\sum_{i=1}^{3}\frac{\partial\Phi}{\partial x^i}\frac{dx^i}{dt} = -Q\frac{d\Phi}{dt} \tag{20}$$

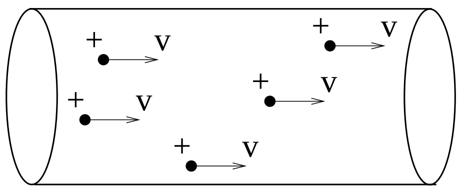
so that

$$\mathcal{E} = \frac{1}{2}m|\vec{v}|^2 + Q\Phi = \text{constant} \tag{21}$$

In the motion of the charged particle, the kinetic energy changes with time, but the sum of the kinetic energy and $Q\Phi$ is constant; thus $Q\Phi$ is the *potential energy* of the particle of charge Q in the electric field with potential Φ. [1]

1.2 Magnetostatics

Frequently, people find magnetism a little more mysterious than electricity. Rather than thinking of fridge magnets, you should think of electromagnets: charges in motion give rise to electric currents and these produce magnetic fields; many of you will have visualized these fields in experiments by sprinkling iron filings on sheets of cardboard transverse to a current-carrying wire. (The fridge magnet, or any permanent magnet, derives its magnetism from microscopic currents.) The subject corresponding to electrostatics, which is produced by time-independent charges, is magnetostatics, which is produced by time-independent, or "steady", currents.



Given a collection of point charges Q_i in motion with velocities $\vec{v}_i$, we define the corresponding electric currents as:

$$\vec{j} = \sum Q_i \vec{v}_i$$

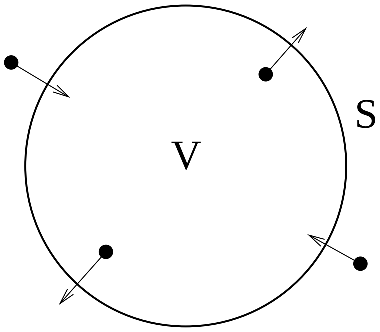
This is a vector field made by adding up elementary contributions. By analogy with charge density ρ as a smoothed-out distribution of point charges, we introduce a vector field, the current density $\vec{J}(x, y, z, t)$, so that the total electric current in a region V is the integral $\int_V \vec{J} dV$.

In this section we shall usually be thinking of steady currents, but there is one thing to deal with first, namely the mathematical expression for the physical observation than charge is conserved.

Conservation of charge

Consider a region V with surface S with charges moving through.

[1] Note that this is completely analogue to the usual (gravitational) potential energy, the mass m is replaced by the charge Q and the potential gh is replaced by the potential Φ.



The total charge inside V is $Q = \int_V \rho dV$, so that

$$
\begin{aligned}
\frac{dQ}{dt} &= \int_V \frac{\partial \rho}{\partial t} dV \qquad \text{allowing } \tfrac{\partial \rho}{\partial t} \neq 0 \text{ for the moment} \\
&= \text{rate of increase of } Q \\
&= \text{rate charge goes in} - \text{rate charge goes out} \\
&= -\int_S \vec{J} \cdot d\mathbf{S} \\
&= -\int_V \nabla \cdot \vec{J} dV \qquad\qquad \text{by divergence theorem}
\end{aligned}
$$

an so

$$\int_V \left(\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{J} \right) dV = 0 \tag{22}$$

this is to be true for all regions V, then

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{J} = 0 \tag{23}$$

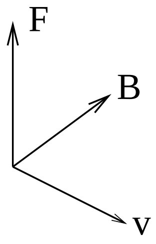
which is the *charge conservation equation.*

For the rest of this section, we suppose none of the quantities of interest depends on time. For this time independent situation (23) reduces to $\nabla \cdot \vec{J} = 0$.

Magnetic Field Strength

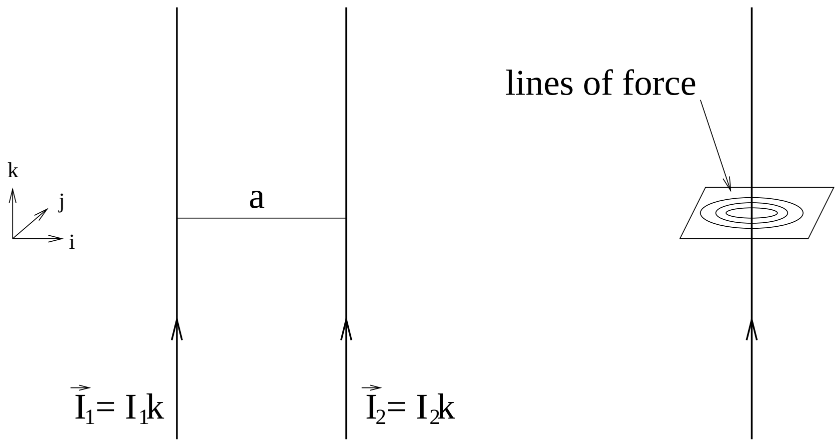
Like the electric field $\vec{E}$, the magnetic field strength $\vec{B}$ can be defined from an experimentally established force law. Analogously to (7), a charged particle with charge Q, moving with velocity $\vec{v}$ in a magnetic field $\vec{B}$ is subject to a force

$$\vec{F} = Q\vec{v} \wedge \vec{B} \tag{24}$$



This serves to define $\vec{B}$, but note a peculiar feature of this force law, that the force is orthogonal to the velocity. In particular, therefore $\vec{v} \cdot \vec{F} = 0$ and the force does no work.

Forces between two parallel current-carrying wires



To see (24) in action, let us consider the force between two parallel current-carrying wires.

What can we say about the magnetic field $\vec{B}$ generated by a single wire? let us choose coordinates such that the wire is along the z direction and let us use cylindrical-polar coordinates. It can be experimentally verified that if we take a charge Q and move it around the wire, the charge is subject to a force. This force vanishes, *i.e.* there is no force, if the charge moves along the θ direction, while there is a force if the charge moves in any other direction. This means that $\vec{B}$ has a component only in the θ direction[2]; using the symmetry in θ and z we can write $\vec{B} = B(R)\vec{e}_\theta$, where R is the distance to the wire.

Now suppose we have parallel wires, wire 1 carrying current I_1 and wire 2 carrying current I_2. Each wire gives rise to a B-field and the field from one exerts a force on the current in the other. Using (24), and thinking of the currents as charged particles in motion, we have:

Force per unit length on wire 2 due to wire 1 is

$$\vec{F}_{12} = \vec{I}_2 \wedge \vec{B}_1.$$

[2]This is consistent with the following fact: if you sprinkle iron filings on a cardboard held at right angle to a current-carrying wire, you can see that the magnetic lines of force are concentric circles, centered on the wire.

where $\vec{B}_1$ means $\vec{B}$ at wire 2 due to current in wire 1, and from the geometry $\vec{I}_2 = I_2\vec{k}$. Similarly, the force on wire 1 due to wire 2 is

$$\vec{F}_{21} = \vec{I}_1 \wedge \vec{B}_2.$$

where $\vec{B}_2$ is $\vec{B}$ at wire 1 due to wire 2, and $\mathbf{I}_1 = I_1\vec{k}$. These forces must be equal in magnitude and opposite, due to Newton's third law:

$$\vec{I}_2 \wedge \vec{B}_1 = -\vec{I}_1 \wedge \vec{B}_2.$$

From what was said about the direction of $\vec{B}$, we have that at wire 2, $\vec{B}_1 = \vec{j}B_1$ in terms of some magnitude B_1, and in wire 1, $\vec{B}_2 = -\vec{j}B_2$ in terms of some magnitude B_2. Therefore

$$|\vec{F}_{12}| = I_2B_1 = |\vec{F}_{21}| = I_1B_2.$$

It is an experimental fact that this magnitude is

$$|\vec{F}| = \frac{\mu_0}{2\pi a}I_1I_2$$

in terms of a constant μ_0 (which, as with ϵ_0 is needed to get the right dimensions).

From this experimental fact we deduce that, at wire 2,

$$\vec{B} = \frac{\mu_0 I_1}{2\pi a}\vec{j}$$

At a general point, in terms of cylindrical polar coordinates (R, θ, z) with the wire along the z-axis this is

$$= \frac{\mu_0 I}{2\pi}\frac{1}{R}\vec{e}_\theta$$

which we can write in cartesian coordinates as

$$\vec{B} = \frac{\mu_0 I}{2\pi}\left(-\frac{y}{R^2}, \frac{x}{R^2}, 0\right), \qquad R^2 = x^2 + y^2 \tag{25}$$

This is the magnetic field due to an infinite straight wire with constant current, which can be though of as the elementary magnetic field, much as (4), $\vec{E} = \frac{Q}{4\pi\epsilon_0}\frac{\vec{r}}{r^3}$ gives the elementary electric field.

From (25), for $R \neq 0$, we calculate

$$\nabla \cdot \vec{B} = 0, \qquad \nabla \wedge \vec{B} = 0 \tag{26}$$

Our aim for the next few pages is to calculate the $\vec{B}$ due to current in an arbitrary wire, much as the integral (18) together with the relation $\vec{E} = -\nabla\Phi$ gives $\vec{E}$ for an arbitrary

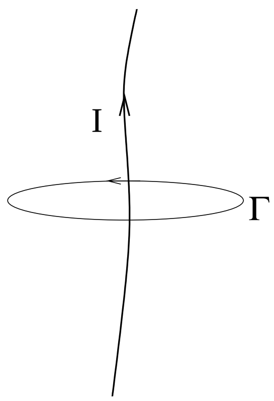
charge distribution. We achieve this aim with (32) below. As a step towards this aim, we note that, when Γ is a horizontal circle of radius R centered on a straight wire

$$\oint_\Gamma \vec{B}\cdot d\vec{\ell} = \frac{\mu_0 I}{2\pi}\oint \frac{1}{R}\vec{e}_\theta\cdot\vec{e}_\theta R d\theta \tag{27}$$

$$= \mu_0 I \tag{28}$$

Even though we have derived this for a straight wire, the result is much more general:

Ampere's Law: $\oint_\Gamma \vec{B}\cdot d\vec{\ell} = \mu_0\times$ total current through Γ.



Compare ampere's law to Gauss Law in the first form we had.

For a current density instead of a wire, we obtain an integral version of Ampere's Law:

$$\oint_\Gamma \vec{B}\cdot d\vec{\ell} = \mu_0\int_\Sigma \vec{J}\cdot d\mathbf{S} \qquad \Sigma \text{ spans } \Gamma$$

$$= \int_\Sigma \nabla\wedge\vec{B}\cdot d\mathbf{S} \qquad \text{by Stoke's theorem}$$

and for this to be true for all Σ we obtain the differential version of Ampere's law:

$$\nabla\wedge\vec{B} = \mu_0\vec{J} \tag{29}$$

(recall that $\nabla\cdot\vec{J} = 0$ from charge conservation and time independence, and we need this for (29) to make sense).

By looking at (26) we have $\nabla\cdot\vec{B} = 0$, which is to be contrasted with (15): $\nabla\cdot\vec{E} = \frac{1}{\epsilon_0}\rho$. Thus we interpret (26) as saying *"there are no magnetic charges"*.

The magnetic potential

It follows from (9) that, in a simply-connected region, there exists a function Φ such that $\vec{E} = -\nabla\Phi$. It is a less familiar fact that

Claim: *if $\nabla \cdot \vec{B} = 0$ in a suitable region, then there exists in that region a vector field $\vec{A}$ such that $\vec{B} = \nabla \wedge \vec{A}$.*

You will prove this on problem sheet 2. $\vec{A}$ is the *magnetic potential* (also called *vector* potential, in which case Φ is called the *scalar* potential).

Note that a change $\vec{A} \to \vec{A} + \nabla\zeta$, for any scalar function ζ, leaves $\vec{B}$ unchanged. We may exploit this freedom to impose another condition, namely

$$\nabla \cdot \vec{A} = 0$$

for suppose $F = \nabla \cdot \vec{A} \neq 0$ and change $\vec{A} \to \vec{A} + \nabla \cdot \zeta$, then this would change $F \to F + \nabla^2\zeta$; now, choose ζ such that $\nabla^2\zeta = -F$, so that now $\nabla \cdot \vec{A} = 0$.

Equation (29) can be turned into an equation for $\vec{A}$ as follows:

$$\begin{aligned} \nabla \wedge \vec{B} &= \nabla \wedge (\nabla \wedge \vec{A}) = \nabla(\nabla \cdot \vec{A}) - \nabla^2 \vec{A} \\ &= \mu_0 \vec{J} \end{aligned} \tag{30}$$

with the choice $\nabla \cdot \vec{A} = 0$, then we have

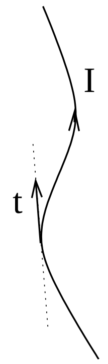
$$\nabla^2 \vec{A} = -\mu_0 \vec{J} \tag{31}$$

Remember that this is true for the particular choice $\nabla \cdot \vec{A} = 0$. This has the vector form of the *Poisson's equation*, and we've solved Poisson's equation before: recall (18) and compare with (16). So we solve (31) by

$$\vec{A} = \frac{\mu_0}{4\pi} \iiint \frac{\vec{J}(\vec{r}', t)}{|\vec{r} - \vec{r}'|} dx' dy' dz'$$

This is for a volume distribution of current. If we have just a line current $\vec{J} = I\vec{t}$, where I is constant and $\vec{t}$ is the tangent (of unit length) to a curve L (where the curve L is the "wire"), then instead

$$\vec{A} = \frac{\mu_0 I}{4\pi} \int_L \frac{\vec{t} d\ell}{|\vec{r} - \vec{r}'(\ell)|}$$



where ℓ parametrizes where we are along the wire.

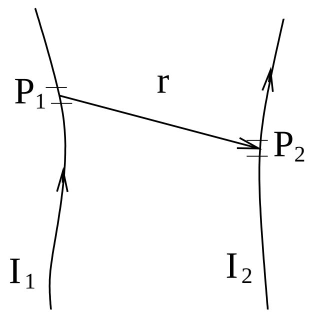
From $\vec{A}$ we calculate $\vec{B}$ as

$$\begin{aligned}\vec{B} &= \nabla \wedge \vec{A} = \frac{\mu_0 I}{4\pi}\int_L \nabla \wedge \left(\frac{\vec{t}}{|\vec{r}-\vec{r}'|}\right) d\ell \\ &= \frac{\mu_0 I}{4\pi}\int_L \frac{\vec{t}\wedge(\vec{r}-\vec{r}')}{|\vec{r}-\vec{r}'|^3} d\ell\end{aligned}$$

or

$$\vec{B} = \frac{\mu_0 I}{4\pi}\int_L \frac{d\vec{\ell}\wedge(\vec{r}-\vec{r}')}{|\vec{r}-\vec{r}'|^3} \tag{32}$$

Where $d\vec{\ell} = d\ell\vec{t}$. This is now $\vec{B}$ at all points of space due to an arbitrary wire, and so generalizes (25). We can use it to obtain an expression for the force between two arbitrary wires, by chopping the second wire into elementary pieces.



At P_2 on wire 2

$$d\vec{F} = I_2 d\vec{\ell}_2 \wedge \vec{B}_1$$

so

$$\begin{aligned}\vec{F} &= \int I_2 d\vec{\ell}_2 \wedge \vec{B}_1 \\ &= \frac{\mu_0 I_1 I_2}{4\pi}\iint \frac{d\vec{\ell}_2 \wedge (d\vec{\ell}_1 \wedge (\vec{r}_2 - \vec{r}_1))}{|\vec{r}_2 - \vec{r}_1|^3}\end{aligned} \tag{33}$$

where $\vec{r}_1 = \vec{r}_1(\ell_1)$ belongs to wire 1 and $\vec{r}_2 = r_2(\ell_2)$ belongs to wire 2. This is the *Biot-Savart Law* for the force between two arbitrary current-carrying wires (more precisely, the force on wire 2 done by wire 1). Its interest is largely theoretical, as it is rather hard to use.

1.3 The story so far

We've studied time-independent electricity (or electrostatics) and time-independent magnetism (or magnetostatics), and found these to be governed by the following system of equations

$$\begin{aligned} \nabla \cdot \vec{E} = \frac{1}{\epsilon_0}\rho \qquad & \nabla \cdot \vec{B} = 0 \\ \nabla \wedge \vec{E} = 0 \qquad & \nabla \wedge \vec{B} = \mu_0 \vec{J} \end{aligned} \tag{34}$$

with the understanding that

$$\frac{\partial \vec{E}}{\partial t} = 0 = \frac{\partial \vec{B}}{\partial t} = \frac{\partial \vec{J}}{\partial t} = \frac{\partial \rho}{\partial t}.$$

The electric and magnetic field strengths themselves can be defined from the force laws (7) and (24), which we combine as follows: a particle with charge Q moving with velocity $\vec{v}$ in a combination of an electric field $\vec{E}$ and a magnetic field $\vec{B}$ is subject to the force given by

$$\vec{F} = Q(\vec{E} + \vec{v} \wedge \vec{B}).$$

In addition we have one equation which we expect to be valid also in the time-dependent case, namely the charge conservation equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{J} = 0. \tag{35}$$

The problem now is how to extend all these equations to the time-dependent case, i.e. how to put back the terms that are zero in the time-independent case. This will be the subject of the second part of this course.

2 Time Dependent Electromagnetism: Maxwell's Equations

Experimental evidence suggests two qualitative facts:

- A time-varying $\vec{B}$ produces an $\vec{E}$.
- A time-varying $\vec{E}$ produces a $\vec{B}$.

This experimental evidence can be summarized quantitatively in the following two integral laws:

$$\oint_\Gamma \vec{E}\cdot d\vec{\ell} = -\frac{d}{dt}\int_\Sigma \vec{B}\cdot d\mathbf{S} \qquad \text{Faraday's law of induction} \tag{36}$$

$$\oint_\Gamma \vec{B}\cdot d\vec{\ell} = \int_\Sigma \mu_0 \vec{J}\cdot d\mathbf{S} + \int_\Sigma \epsilon_0\mu_0 \frac{\partial \vec{E}}{\partial t}\cdot d\mathbf{S} \tag{37}$$

Where Σ spans Γ. Apply Stoke's theorem to both to obtain

$$\int_\Sigma \left(\nabla\wedge\vec{E} + \frac{\partial \vec{B}}{\partial t}\right)\cdot d\mathbf{S} = 0 \tag{38}$$

$$\int_\Sigma \left(\nabla\wedge\vec{B} - \mu_0\vec{J} - \epsilon_0\mu_0\frac{\partial \vec{E}}{\partial t}\right)\cdot d\mathbf{S} = 0 \tag{39}$$

and if these hold for all Σ we obtain the

Maxwell equations:

$$\begin{aligned} \nabla\wedge\vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0; & \nabla\wedge\vec{B} &= \mu_0\vec{J} + \frac{1}{c^2}\frac{\partial \vec{E}}{\partial t} \\ \nabla\cdot\vec{B} &= 0, & \nabla\cdot\vec{E} &= \frac{1}{\epsilon_0}\rho \end{aligned} \tag{40}$$

with $c^2 = \frac{1}{\epsilon_0\mu_0}$.

These are the *Maxwell's equations*, the basic equations of the new theory of electromagnetism, in which electricity and magnetism are merged.

Note that

- Two of the equations always have zero on the right; the other two contain the sources ρ and $\vec{J}$.

- If all time-derivatives are set to zero, we recover the equations of the previous section.
- The system obtained by setting ρ and $\vec{J}$ to zero is called *source-free Maxwell equations*.
- Regarding ρ and $\vec{J}$ as given, there are eight equations for six unknowns, so there should be a consistency condition.

To see what the consistency condition is, calculate:

$$\begin{aligned}
\frac{\partial \rho}{\partial t} &= \epsilon_0 \frac{\partial}{\partial t}(\nabla \cdot \vec{E}) = \epsilon_0 \nabla \cdot \left(\frac{\partial \vec{E}}{\partial t} \right) \\
&= \frac{1}{\mu_0} \nabla \cdot \left(\nabla \wedge \vec{B} - \mu_0 \vec{J} \right) \\
&= -\nabla \cdot \vec{J}
\end{aligned}$$

which we recognize as the charge conservation equation. From this point of view the charge conservation equation is a consistency condition for the Maxwell's equations. As an exercise, show that there is a second consistency condition but that it is automatically satisfied.

Energy of the Electromagnetic Field

There is a deductive way to approach this question, but we shall rely on intuition and guesswork. We want an energy density (energy-per-unit-volume) which is something like "$1/2m\vec{v} \cdot \vec{v}$".

After playing around with Maxwell equations for a while, we are lead to consider

$$\mathcal{E} = \frac{1}{2}\epsilon_0 |\vec{E}|^2 + \frac{1}{2\mu_0} |\vec{B}|^2 \tag{41}$$

Then, using the Maxwell equations, we calculate

$$\begin{aligned}
\frac{\partial \mathcal{E}}{\partial t} &= \epsilon_0 \vec{E} \cdot \frac{\partial \vec{E}}{\partial t} + \frac{1}{\mu_0} \vec{B} \cdot \frac{\partial \vec{B}}{\partial t} \\
&= \frac{1}{\mu_0} \vec{E} \cdot \left(\nabla \wedge \vec{B} - \mu_0 \vec{J} \right) - \frac{1}{\mu_0} \vec{B} \cdot \nabla \wedge \vec{E} \\
&= -\nabla \cdot \left(\frac{1}{\mu_0} \vec{E} \wedge \vec{B} \right) - \vec{E} \cdot \vec{J}
\end{aligned}$$

For the source-free case, we set $\vec{J} = 0$ so that the last term vanishes. Then, this has the form of a conservation equations

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot \vec{P} = 0 \tag{42}$$

Remember that $\mathcal{E}$ is the energy density. $\vec{P} \equiv \frac{1}{\mu_0}\vec{E} \wedge \vec{B}$ is called the *Poynting vector* and has the interpretation of the rate of energy flow, or momentum density.

If we now reinstate the sources, this equation has a name:

Poynting's theorem

$$\frac{\partial \mathcal{E}}{\partial t} = -\nabla \cdot \vec{P} - \vec{E} \cdot \vec{J}$$

We integrate Poynting's theorem over a volume V with surface S to obtain an energy balance equation:

$$\begin{array}{ccccc} \dfrac{d}{dt}\displaystyle\int_V \mathcal{E} dV & = & -\displaystyle\int_S \vec{P} \cdot d\mathbf{S} & - & \displaystyle\int_V \vec{E} \cdot \vec{J} dV \\ (1) & & (2) & & (3) \end{array}$$

where each term has the following interpretation:

1. Rate of increase of electromagnetic energy.
2. Rate of energy flow into V.
3. Rate of work done by the field on sources.

To justify the interpretation of (3), consider a single charge. Remember that work is the product of force and displacement $dW = \vec{F} \cdot d\vec{\ell}$, so the rate of work is $\frac{dW}{dt} = \vec{F} \cdot \vec{v}$. For a single charge $\vec{F} = Q\vec{E}$ and $Q\vec{v}$ is by definition the current $\vec{j}$, so that $\frac{dW}{dt} = \vec{E} \cdot \vec{j}$. This satisfactory interpretation of the third term reinforces the interpretation of $\mathcal{E}$ as energy density.

Potentials

The introduction of potentials in the time-dependent case is similar to that in the time-independent case. Assume we are interested in a suitable (simply-connected, etc) region of space, then the Maxwell equations imply, first:

$$\nabla \cdot \vec{B} = 0 \Rightarrow \exists \vec{A} \text{ such that } \vec{B} = \nabla \wedge \vec{A};$$

next $\frac{\partial \vec{B}}{\partial t} = \nabla \wedge \frac{\partial \vec{A}}{\partial t}$, so that

$$\nabla \wedge \vec{E} + \frac{\partial \vec{B}}{\partial t} = \nabla \wedge \left(\vec{E} + \frac{\partial \vec{A}}{\partial t} \right) = 0,$$

whence

$$\exists \Phi \text{ such that } \vec{E} + \frac{\partial \vec{A}}{\partial t} = -\nabla \Phi$$

i.e.

$$\vec{B} = \nabla \wedge \vec{A} \tag{43}$$

$$\vec{E} = -\nabla\Phi - \frac{\partial \vec{A}}{\partial t} \tag{44}$$

The freedom to modify $\vec{A}$ and Φ changes slightly: if $\vec{A} \to \vec{A} + \nabla\zeta$, then $\vec{B}$ is unchanged, but we need $\Phi \to \Phi - \frac{\partial \zeta}{\partial t}$ to leave $\vec{E}$ unchanged.

This is a *gauge transformation* and can be exploited, as before, to simplify potentials.

Electromagnetic Waves

For the rest of this section, we show that the source-free Maxwell equations admit wave solutions, and use these solutions to recover some elementary properties of optics. The realization that light is an electromagnetic phenomenon was a triumph of nineteenth century science.

We start from the Source-free Maxwell equations:

$$\begin{aligned} \nabla \cdot \vec{B} &= 0, & \nabla \cdot \vec{E} &= 0 \\ \nabla \wedge \vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0, & \nabla \wedge \vec{B} - \tfrac{1}{c^2}\tfrac{\partial \vec{E}}{\partial t} &= 0 \end{aligned} \tag{45}$$

we would like to decouple the equations for $\vec{E}$ and $\vec{B}$. Calculate:

$$\begin{aligned} \nabla \wedge (\nabla \wedge \vec{E}) &= \nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E} = -\nabla^2 \vec{E} \\ &= -\nabla \wedge \frac{\partial \vec{B}}{\partial t} = -\frac{\partial}{\partial t}(\nabla \wedge \vec{B}) = -\frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} \end{aligned}$$

so that

$$\frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} = \nabla^2 \vec{E} \tag{46}$$

which is the *wave-equation* with $c = (\epsilon_0 \mu_0)^{-1/2}$ as the wave speed.

Similarly (exercise)

$$\frac{1}{c^2}\frac{\partial^2 \vec{B}}{\partial t^2} = \nabla^2 \vec{B} \tag{47}$$

We want to solve these wave equations, so think first about a scalar version

$$\frac{1}{c^2}\frac{\partial^2 F}{\partial t^2} = \nabla^2 F = \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} + \frac{\partial^2 F}{\partial z^2} \tag{48}$$

and try $F = f(\omega t - \vec{K}\cdot\vec{x})$ with constant ω and $\vec{K}$. Then

$$\frac{\partial F}{\partial t} = \omega f', \qquad \nabla F = -\vec{K} f'$$
$$\frac{\partial^2 F}{\partial t^2} = \omega^2 f'', \qquad \nabla^2 F = |\vec{K}|^2 f''$$

Hence, $F = f(\omega t - \vec{K}\cdot\vec{x})$ is a solution for any (twice-differentiable) f provided $|\vec{K}|^2 = \frac{\omega^2}{c^2}$. Note that F is constant on surfaces

$$\omega t - \vec{K}\cdot\vec{x} = \text{const}$$

at a fixed t this is the equation of a plane, so these solutions are called *planewaves*. If f is cos or sin, then they are called *harmonic waves* with a single frequency, or *monochromatic plane waves*.

Now try

$$\begin{aligned} \vec{E} &= \vec{e}_1\cos\phi + \vec{e}_2\sin\phi, \qquad \phi = \omega t - \vec{K}\cdot\vec{x} \\ \vec{B} &= \vec{b}_1\cos\phi + \vec{b}_2\sin\phi, \end{aligned} \tag{49}$$

with constant vectors $\vec{K}$ $\vec{e}_i$ and $\vec{b}_i$. Imposing the source-free Maxwell equations in order:

$$\begin{aligned} \nabla\cdot\vec{E} &= -\sin\phi\vec{e}_1\cdot\nabla\phi + \cos\phi\vec{e}_2\cdot\nabla\phi \\ &= \sin\phi\vec{K}\cdot\vec{e}_1 - \cos\phi\vec{K}\cdot\vec{e}_2 \\ &= 0 \end{aligned}$$

so we need

$$\vec{K}\cdot\vec{e}_1 = \vec{K}\cdot\vec{e}_2 = 0 \;\text{ and then }\; \vec{K}\cdot\vec{E} = 0 \tag{50}$$

In the same way we also have

$$\nabla\cdot\vec{B} = 0 \Rightarrow \vec{K}\cdot\vec{b}_1 = \vec{K}\cdot\vec{b}_2 = 0 \;\text{ and then }\; \vec{K}\cdot\vec{B} = 0 \tag{51}$$

Next

$$\nabla\wedge\vec{E} + \frac{\partial\vec{B}}{\partial t} = 0$$

so

$$\vec{K}\wedge\vec{e}_1\sin\phi - \vec{K}\wedge\vec{e}_2\cos\phi - \omega\vec{b}_1\sin\phi + \omega\vec{b}_2\cos\phi = 0$$

i.e.

$$\omega\vec{b}_i = \vec{K}\wedge\vec{e}_i \tag{52}$$

and so

$$\omega\vec{B} = \vec{K} \wedge \vec{E} \tag{53}$$

Finally

$$\nabla \wedge \vec{B} - \frac{1}{c^2}\frac{\partial \vec{E}}{\partial t} = 0$$

so

$$\vec{K} \wedge \vec{b}_1 \sin\phi - \vec{K} \wedge \vec{b}_2 \cos\phi + \frac{\omega}{c^2}\vec{e}_1 \sin\phi - \frac{\omega}{c^2}\vec{e}_2 \cos\phi$$

or

$$\frac{\omega}{c^2}\vec{e}_i = -\vec{K} \wedge \vec{b}_i \tag{54}$$

$$\frac{\omega}{c^2}\vec{E} = -\vec{K} \wedge \vec{B} \tag{55}$$

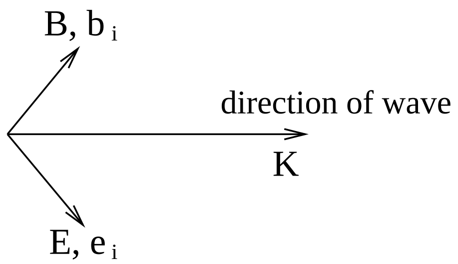
From this and (52) we deduce:

$$\begin{aligned}
\frac{\omega^2}{c^2}\vec{e}_i &= \omega\vec{b}_i \wedge \vec{K} \\
&= -(\vec{e}_1 \wedge \vec{K}) \wedge \vec{K} \\
&= |\vec{K}|^2 \vec{e}_i
\end{aligned}$$

so that $|\vec{K}|^2 = \frac{\omega^2}{c^2}$ as expected and (52) implies $\omega|\vec{b}_i| = |\vec{e}_i||\vec{K}|,\ \vec{e}_i \cdot \vec{b}_i = 0$, so

$$|\vec{e}_i| = c|\vec{b}_i|, \quad \vec{e}_i \cdot \vec{b}_i = 0 \tag{56}$$

These are *transverse waves* (unlike *e.g.* sound). The waves travel in the direction of $\vec{K}$, while $\vec{E}$ and $\vec{B}$ lie in wave fronts, transverse to the direction of propagation, and are orthogonal to each other.



To simplify the expressions, take $\vec{K}$ in the z-direction, $\vec{K} \cdot \vec{x} = Kz$. Look in the plane $z = 0$ so that $\phi = \omega t$. Then

$$\vec{E} = \vec{e}_1 \cos\omega t + \vec{e}_2 \sin\omega t \tag{57}$$

$$\vec{B} = \frac{1}{\omega}\vec{K} \wedge \vec{e}_1 \cos\omega t + \frac{1}{\omega}\vec{K} \wedge \vec{e}_2 \sin\omega t \tag{58}$$

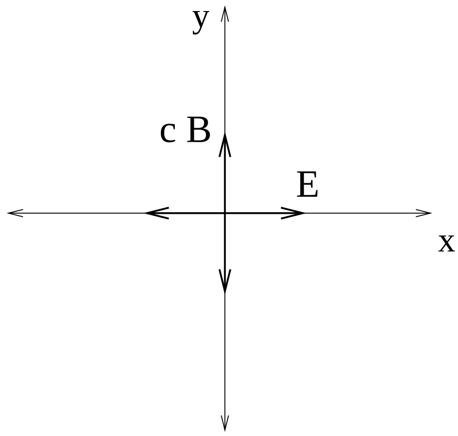
To discuss the phenomenon of *polarisation*, we distinguish some particular cases:

(i) $\vec{e}_1, \vec{e}_2$ proportional, say to $\vec{i}$, then

$$\vec{E} = E\vec{i}\cos(\omega t + \delta) \tag{59}$$

$$\vec{B} = \frac{1}{c}E\vec{j}\cos(\omega t + \delta) \tag{60}$$

for some constant δ;



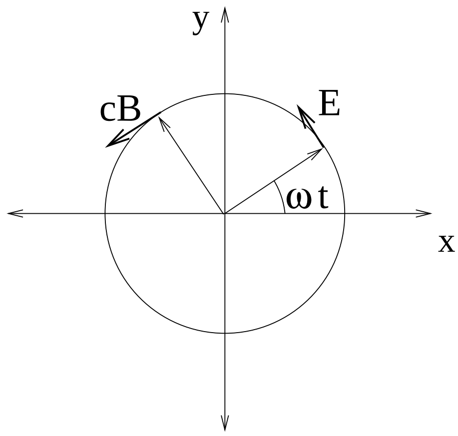
a wave like this is said to be *lineary polarised*; $\vec{E}$ and $\vec{B}$ oscillate parallel to two fixed orthogonal directions.

(ii) $\vec{e}_1, \vec{e}_2$ orthogonal, of equal length, e.g. $\vec{e}_1 = E\vec{i}, \vec{e}_2 = E\vec{j}$, then

$$\vec{E} = E(\vec{i}\cos\omega t + \vec{j}\sin\omega t) \tag{61}$$

$$\vec{B} = \frac{E}{c}(-\vec{i}\sin\omega t + \vec{j}\cos\omega t); \tag{62}$$

a wave like this is said to be *circularly polarised*; $\vec{E}$ and $\vec{B}$ rotate at a constant rate about the direction of propagation.



The direction is anti-clockwise and the wave is *left circularly polarized* if $\vec{e}_1 \wedge \vec{e}_2 \cdot \vec{K} > 0$, and the rotation is clockwise and the wave is *right circularly polarized* if $\vec{e}_1 \wedge \vec{e}_2 \cdot \vec{K} < 0$.

(iii) the general case can be regarded as a combination of waves with different polarizations.

Epilogue

What comes next in this story? The question *how does one transform the Maxwell equations between frames of reference which are moving uniformly relative to each other?* Evidently, waves in one frame should be waves in the other, but the big surprise is the experimental fact that the *speed of light in vacuum is the same for all observers in uniform relative motion!*. Since the speed of light is made from constants appearing in Maxwell's equations, this at least suggests that Maxwell's equations are the same in all frames. How all this works is the province of Special Relativity: first time and space fit together as space-time, then t and $\vec{r}$ fit together so that vectors have four components instead of three, now Φ and $\vec{A}$ form a 4-vector and so do ρ and $\vec{J}$, while $\vec{E}$ and $\vec{B}$ fit together as an anti-symmetric 4×4 matrix. See B7.2b for more...

Part 2: Quantum Mechanics

Synopsis

The mathematical structure of quantum mechanics and the postulates of quantum mechanics. Commutation relations. Poisson's brackets and Dirac's quantization scheme. Heisenberg's uncertainty principle. Creation and annihilation operators for the harmonic oscillator. Measurements and the interpretation of quantum mechanics. Schroedinger's cat. Spin 1/2 *particles. Angular momentum in quantum mechanics. Particle in a central potential. The hydrogen atom.*

Reading list

The main book followed for this part of the course is:

K. C. Hannabuss, Introduction to Quantum Mechanics (Oxford University Press, 1997).

Classic texts on Quantum Mechanics:

L. I. Schiff, Quantum Mechanics (3rd edition, Mc Graw Hill, 1968)

L D Landau and E.M. Lifshitz, Quantum Mechanics Non-Relativistic Theory: Volume 3.

3 The Mathematical Structure of Quantum Mechanics

(Hannabuss chapter 6)

3.1 Reviewing Part A Course

We shall recall what you learned there and add one new idea:

- a particle is described by its *wave function*:

$$\begin{aligned}\psi(\vec{x},t) &= \psi(x,y,z,t) \;\text{ in 3 dim} \\ &= \psi(x,t) \;\text{ in 1 dim}\end{aligned}$$

  which is complex and continuous (and usually twice-differentiable in $\vec{x}$).

- the function $\rho \equiv |\psi|^2$ is interpreted as the *probability density of position* provided that

$$\int \rho = \int |\psi|^2 = 1 \tag{63}$$

  i.e. provided ψ is *normalized*. (We'll usually omit the range of integration when its intended to be all space). For this interpretation to make sense we at least need

$$\int |\psi|^2 < \infty, \quad \textit{i.e. } \psi \text{ is normalizable,} \tag{64}$$

  for then a constant multiple of ψ is normalized. Recall also the definition of the *probability current*.

$$\vec{j} \equiv -\frac{i\hbar}{2m}(\bar{\psi}\nabla\psi - \psi\nabla\bar{\psi}) \tag{65}$$

- for a particle of mass m moving in a real potential $V(\vec{x})$ ψ changes in time according to the *Schrodinger equation*:

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi \equiv \left(-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{x})\right)\psi \tag{66}$$

  This is motivated by the Hamiltonian from Classical Mechanics

$$H = \frac{1}{2m}|\vec{p}|^2 + V = E \tag{67}$$

  which gives the total energy E, together with the substitution

$$\vec{p} \to -i\hbar\nabla. \tag{68}$$

- the Schrodinger equation is linear, so the solutions of (66) form a complex vector space, and given a normalizable solution, we can normalize it (multiplying it by an appropriate constant) and still have a solution.

- there exist separable solutions of (66):

$$\psi(\vec{x}, t) = \Psi(\vec{x})T(t) \tag{69}$$

with

$$T(t) = e^{-\frac{iEt}{\hbar}} \tag{70}$$

and

$$H\Psi = E\Psi \tag{71}$$

which is an eigenvalue equation for the allowed energies. Solutions like (69) are called *stationary states* even though $\frac{\partial\psi}{\partial t} \neq 0$, because $\frac{\partial\rho}{\partial t} = 0 = \vec{j}$. Correspondingly (71) is called the *stationary state Schrodinger equation*.

- in many examples (e.g. particle in a box, harmonic oscillator,...) (71) has a countably infinite, orthonormalizabled set of solutions $\{\Psi_n : n \in \mathbb{N}\}$ with corresponding real eigenvalues E_n, so that, without loss of generality

$$\int \bar{\Psi}_i \Psi_j = \delta_{ij}. \tag{72}$$

These provide a basis for the general solution of (66):

$$\psi(\vec{x}, t) = \sum_k c_k \Psi_k(\vec{x}) e^{-iE_k t/\hbar} \tag{73}$$

If we know the wave function at the initial time $\psi(\vec{x}, 0)$, then we can determine the constants c_k:

$$c_k = \int \overline{\Psi}_k(\vec{x})\psi(\vec{x}, 0) \tag{74}$$

Hence, the wave function at initial time determines the wave function at all times (as it should, since the Schrodinger equation is a first order equation in t). Now the normalization of ψ becomes

$$1 = \int |\psi|^2 = \int \left(\sum \bar{c}_j \overline{\Psi}_j e^{iE_j t/\hbar}\right)\left(\sum c_k \Psi_k e^{iE_k t/\hbar}\right) = \sum |c_k|^2$$

- **NEW IDEA:** Now comes the new idea: since they sum to one, it is tempting to think of the c_j as the probabilities that the particle is in state j, and we shall. Since the state j has energy E_j this implies

$$P(E = E_j) = |c_j|^2,$$

writing P for probability, and then the average value, or more properly the expectation, of the energy is

$$\mathbb{E}(H) = \sum_j E_j |c_j|^2, \tag{75}$$

note that this coincides with the usual definition of expectation

$$\begin{aligned} \langle H \rangle_\psi = \int \bar{\psi} H \psi &= \int \bar{\psi} i\hbar \frac{\partial \psi}{\partial t} \\ &= \int \left(\sum \bar{c}_j \bar{\Psi}_j e^{iE_j t/\hbar} \right) \left(\sum c_k E_k \Psi_k e^{-iE_k t/\hbar} \right) \\ &= \sum E_j |c_j|^2 \end{aligned}$$

where we have used the definition from the Part A course for the *expectation of a function f of position ($\vec{x}$) in the state ψ*:

$$\mathbb{E}(f) \equiv \langle f(\vec{x}) \rangle_\psi = \int f(\vec{x}) |\psi(\vec{x}, t)|^2 d\vec{x} = \int \bar{\psi} f \psi \tag{76}$$

3.2 States and Observables

Having finished the review, we start this part of the course by seeking a more abstract setting for the above. We suppose that, in quantum mechanics,

- a system can take up one of a number of *states* which correspond to elements of a complex vector space $\mathcal{H}$, called the Hilbert space, which has a Hermitian inner product that we will write $\langle u|v \rangle$ for $u, v \in \mathcal{H}$
- proportional vectors define the same state (so states actually correspond to one-dimensional subspaces or *rays* in $\mathcal{H}$).
- an element $\psi \in \mathcal{H}$ is normalized if

  $$||\psi||^2 \equiv \langle \psi | \psi \rangle = 1$$

  (this still leaves a freedom $\psi \to \psi e^{i\alpha}$ without changing the state.)
- observables are self-adjoint linear operators $A : \mathcal{H} \to \mathcal{H}$ (recall the adjoint A^* of a linear operator A is defined by

  $$\langle A^* \phi | \psi \rangle = \langle \phi | A \psi \rangle, \quad \text{for all } \phi, \psi \in \mathcal{H}$$

  and A is *self-adjoint* if $A^* = A$)

- the mot important example of an $\mathcal{H}$ is complex-valued functions ψ on $\mathbb{R}^3$ with

$$\langle\phi|\psi\rangle \equiv \int_{\mathbb{R}^3} \bar{\phi}\psi \tag{77}$$

then the *position operator* $\vec{X} = \{X_i, i = 1, 2, 3\}$ and the *momentum operator* $\vec{P} = \{P_i, i = 1, 2, 3\}$ are defined (as we know) by

$$\begin{aligned} X_j\psi &= x_j\psi \\ P_j\psi &= -i\hbar\frac{\partial\psi}{\partial x_j} \end{aligned} \tag{78}$$

Are these in fact self-adjoint? For P_j we calculate

$$\langle\phi|P_j\psi\rangle = \int \bar{\phi}(-i\hbar\frac{\partial\psi}{\partial x_j}) = \int (i\hbar\frac{\partial\bar{\phi}}{\partial x_j})\psi = \int \overline{\left(-i\hbar\frac{\partial\phi}{\partial x_j}\right)}\psi = \langle P_j\phi|\psi\rangle$$

on integration by parts, assuming differentiability of ϕ, ψ and the vanishing of the integrated term, which would follow from normalizability of ϕ, ψ. Thus the P_j are self-adjoint, and it is trivial to check that the X_j are. (Note however, that the P_j are only defined on a subspace of $\mathcal{H}$ (normalizable functions), and that both X_i and P_i may spoil normalizability- we will note but pass over this kind of technical issue.)

3.3 Statistical Issues

Following section 3.1, we will define the *expectation* of an observable A in a (normalizable) state ψ to be

$$\mathbb{E}_\psi(A) \equiv \frac{\langle\psi|A\psi\rangle}{||\psi||^2} \tag{79}$$

then e.g. for a function $F(\vec{X})$ and a normalized ψ,

$$\mathbb{E}_\psi(F) = \langle\psi|F\psi\rangle = \int \bar{\psi}F\psi \tag{80}$$

just as before. From the abstract definition (79) we obtain

1. $\mathbb{E}_\psi(I) = 1$ for the identity operator I;
2. $\mathbb{E}_\psi(A)$ is real for self-adjoint A;
3. $\mathbb{E}_\psi(A) \geq 0$ for positive A (i.e. A for which $\langle\psi|A\psi\rangle \geq 0$ for all ψ);
4. $\mathbb{E}_\psi(\alpha A + \beta B) = \alpha\mathbb{E}_\psi(A) + \beta\mathbb{E}_\psi(B)$ for all $\alpha, \beta \,\epsilon\, \mathbb{C}$

All but the second are easy to see; for the second note that

$$\overline{\langle\psi|A\psi\rangle} = \langle A\psi|\psi\rangle = \langle\psi|A^*\psi\rangle = \langle\psi|A\psi\rangle$$

for self-adjoint A.

Next define the *dispersion* $\Delta_\psi(A)$ of an observable A in a state ψ as

$$\Delta_\psi(A) = \left(\mathbb{E}_\psi(A^2) - (\mathbb{E}_\psi(A))^2\right)^{1/2} \tag{81}$$

This is like standard deviation, so it measures the spread of the measurements of A about the average value.

Proposition

$\Delta_\psi(A) = 0$ iff ψ is an eigenvector of A.

Proof

Suppose ψ is normalized, then for any real μ,

$$\mathbb{E}_\psi((A-\mu I)^2) = \langle\psi|(A-\mu I)^2\psi\rangle = \langle(A-\mu I)\psi|(A-\mu I)\psi\rangle$$

since $A - \mu I$ is self-adjoint

$$= ||(A-\mu I)\psi||^2$$

Also, by linearity

$$\mathbb{E}_\psi((A-\mu I)^2) = \mathbb{E}_\psi(A^2) - 2\mu\mathbb{E}_\psi(A) + \mu^2$$

so put $\mu = \mathbb{E}_\psi(A)$ (which is allowed, since this is real) to find

$$(\Delta_\psi(A))^2 = ||(A-\mu I)\psi||^2, \qquad \mu = \mathbb{E}_\psi(A)$$

Thus

$$\Delta_\psi(A) = 0 \quad \text{iif} \quad (A-\mu I)\psi = 0 \quad \text{iif} \quad A\psi = \mu\psi$$

as required.

To make contact with earlier formulae, suppose $\mathcal{H}$ has a (countable) basis of normalized eigenvalues of A, say $\{\psi_n\}$, with

$$A\psi_n = \alpha_n\psi_n,$$

then for any normalized $\psi \,\epsilon\, \mathcal{H}$ we have an expansion $\psi = \sum c_n\psi_n$ with $c_n = \langle\psi_n|\psi\rangle$, so that

$$\mathbb{E}_\psi(A) = \langle\psi|A\psi\rangle = \sum_n c_n\langle\psi|A\psi_n\rangle = \sum_n \alpha_n c_n\langle\psi|\psi_n\rangle = \sum_n \alpha_n|c_n|^2,$$

and we recover (75).

Observables are going to be the things we can measure, and the results of these measurements are the eigenvalues of the corresponding linear operator, but we can't assume that every observable of interest has normalizable eigenvectors. For example, consider $\vec{P}$, then the eigenvalue equation is

$$\vec{P}\psi \equiv -i\hbar\nabla\psi = \vec{k}\psi$$

for a constant vector $\vec{k}$, so the eigenfunction is const $\times\, e^{i\vec{k}\cdot\vec{x}/\hbar}$, which isn't normalizable. As another example, suppose

$$\vec{X}\psi = \vec{a}\psi$$

for eigenvalue a. Then

$$(\vec{x}-\vec{a})\psi(\vec{x}) = 0$$

and ψ vanishes at every point except $\vec{x} = \vec{a}$. Continuity would then force ψ to vanish everywhere (though if you are Dirac, you might invent the delta function at this point).

As an example of dispersion, we revisit the treatment of a particle in a box from Part A. If the box occupies $0 \leq x \leq a$, then the Hamiltonian and normalized eigenfunctions were

$$H = \frac{1}{2m}P^2, \qquad \psi_n(x) = \sqrt{\frac{2}{a}}\sin\frac{n\pi x}{a}$$

Now calculate

$$\mathbb{E}_{\psi_n}(P) = \int_0^a \bar{\psi}_n(x)(-i\hbar\psi_n'(x))dx = 0$$

writing prime for d/dx, so that 'on average' the particle moves neither left nor right, while

$$\mathbb{E}_{\psi_n}(P^2) = \mathbb{E}_{\psi_n}(2mH) = 2mE_n = \frac{n^2\pi^2\hbar^2}{a^2}$$

Thus the dispersion of P is not zero - the particle is jiggling about.

3.4 Time Evolution

In this abstract setting we will write ψ_t to indicate a time-dependent element of $\mathcal{H}$, then the time-evolution, following the Schrodinger equation, is determined by a Hamiltonian H, which is an observable, by

$$i\hbar\frac{\partial\psi_t}{\partial t} = H\psi_t \tag{82}$$

Formally this can be solved by the expression[3]

$$\psi_t = \exp\left(-\frac{i}{\hbar}\int_0^t H dt\right)\psi_0 \tag{83}$$

where exp of an operator is defined by

$$\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!}A^k$$

and we will just note that there may be convergence issues.
To prove (83), suppose for simplicity that H has no t-dependence so that

$$\int_0^t H dt = Ht$$

Then introduce

$$U_t \equiv \exp\left(-\frac{iHt}{\hbar}\right) \tag{84}$$

so that (83) is just

$$\psi_t = U_t\psi_0 \tag{85}$$

Claim:

- $U_0 = I$
- $\frac{dU_t}{dt} = -\frac{i}{\hbar}HU_t = -\frac{i}{\hbar}U_tH,$
- $U_sU_t = U_{s+t}$ so in particular $U_{-t} = (U_t)^{-1}$.
- Since H is self-adjoint,

  $$U_t^* = \exp(\frac{iH^*t}{\hbar}) = U_{-t} = (U_t)^{-1}$$

  and so U_t is *unitary*.

The first is clear; the second follows from differentiating (84); for the third, consider

$$W_t = U_sU_t - U_{s+t},$$

[3]The Hamiltonian H could depend on time, but for the following expression to be true we will assume that Hamiltonians at different times always commute. This will always be true for the examples seen in this course.

then $W_0 = 0$ and

$$\frac{dW_t}{dt} = -\frac{i}{\hbar}H(U_sU_t - U_{s+t}) = -\frac{i}{\hbar}HW_t$$

this is a first-order linear ODE so has a unique solution, which must therefore be $W_t = 0$. Then the fourth is clear. Now (85) is equivalent to

$$U_{-t}\psi_t = \psi_0 \tag{86}$$

and we will prove it in this form. Differentiating (86) we obtain

$$\frac{d}{dt}U_{-t}\psi_t = \frac{d}{dt}\left(\exp(\frac{iHt}{\hbar})\psi_t\right) = \exp(\frac{iHt}{\hbar})(\frac{d\psi_t}{dt} + \frac{i}{\hbar}H\psi_t)$$

which vanishes by (82). So $U_{-t}\psi_t$ is independent of t and we can evaluate it at$t = 0$, obtaining (86).

Thus time-evolution following from the Schrodinger equation (82) is a sequence of unitary transformation (85). In particular

$$||\psi_t||^2 = \langle\psi_t|\psi_t\rangle = \langle U_t\psi_0|U_t\psi_0\rangle = \langle\psi_0|U_t^*U_t\psi_0\rangle = \langle\psi_0|\psi_0\rangle = ||\psi_0||^2 \tag{87}$$

i.e. unitary evolution preserves the norm.

3.5 Measurements

When we measure an observable A the result is one of its eigenvalues with some probability. If we measure it again at once, then we should get the same answer with probability one. Thus the dispersion is zero, and the state just after a measurement must be an eigenstate of A with the measured eigenvalue. There may be more than one eigenstate with such eigenvalue, so we are led to make **Luder's Postulate:**

Suppose a measurement of A in a state ψ gives the result α; let Q_α be the projection onto the subspace of $\mathcal{H}$ spanned by the eigenvectors of A with eigenvalues α; then immediately after the measurement ψ has become $Q_\alpha\psi$.

(Recall that a projection is a linear operator with $Q^2 = Q$). We say the measurement has *collapsed* the wave function. The norm changes in this collapse so that it is not unitary.

4 The Commutation Relations

(Hannabuss chapter 7)

4.1 The Commutation Relations

Recall, in one dimension, the definitions of position and momentum operators as

$$X\psi = x\psi, \qquad P\psi = -i\hbar\psi'$$

then $PX\psi = -i\hbar(x\psi)'$ while $XP\psi = x(-i\hbar\psi')$ so that

$$(PX - XP)\psi = -i\hbar\psi \tag{88}$$

Generally, we define the commutator of observables A, B as

$$[A, B] \equiv AB - BA, \tag{89}$$

then we have

$$[P, X] = -i\hbar I, \tag{90}$$

and more generally, in three-dimensions

$$[P_i, X_j] = -i\hbar\delta_{ij} I, \tag{91}$$

For completeness, note also

$$[X_i, X_j] = [P_i, P_j] = 0 \tag{92}$$

Commutators have an algebra which includes the following properties:

1. $[A, B] = -[B, A]$
2. $[A, BC] = [A, B]C + B[A, C]$
3. linearity in A and B, e.g. $[A, \alpha B + \beta C] = \alpha[A, B] + \beta[A, C]$
4. The *Jacobi identity*: $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$.
5. For self-adjoint A, B , $[A, B]$ is not self-adjoint, but $i[A, B]$ is.

I leave the proofs as an exercise!

4.1.1 Aside: Similarity with Poisson Brackets

(Read this only if you know what a Poisson Bracket is)

The commutation relations between X_i and P_i strongly resemble the classical Poisson Bracket relations for generalized coordinates and momenta. Recalling that the Poisson bracket is defined by

$$\{f, g\} = \sum_j \left(\frac{\partial f}{\partial p_j} \frac{\partial g}{\partial x_j} - \frac{\partial f}{\partial x_j} \frac{\partial g}{\partial p_j} \right) \tag{93}$$

we have

$$\{p_i, x_j\} = \delta_{ij}, \qquad \{x_i, x_j\} = \{p_i, p_j\} = 0 \tag{94}$$

Furthermore, the above properties of the commutator also paralell those of the Poisson bracket.

The similarity with Poisson brackets led Dirac to suggest that each function classical f (for instance, the coordinate x or the momentum p) should be related in quantum theory by an operator $Q(f)$, in such a way that for any pair of functions f, g the corresponding operators satisfy

$$[Q(f), Q(g)] = -i\hbar Q(\{f, g\}) \tag{95}$$

This idea doesn't work simultaneously for all operators, and sometimes one faces ambiguities. For instance, to the classical function px^2 we could assign PX^2 or X^2P or PXP, and so on. These differ from each other, since P and X do not commute, but their difference is "small", namely, proportional to $\hbar$.

4.2 Heisenberg's Uncertainty Principle

We will start with a

Lemma

Suppose A, B, C are self-adjoint with

$$[A, B] = iC$$

and μ is real. Then

1. $(A - i\mu B)^*(A - i\mu B) = A^2 + \mu C + \mu^2 B^2$;
2. for normalized $\psi \in \mathcal{H}$
$$||(A - i\mu B)\psi||^2 = \mathbb{E}_\psi(A^2) + \mu\mathbb{E}_\psi(C) + \mu^2\mathbb{E}_\psi(B^2);$$
3. for any normalizable $\psi \in \mathcal{H}$
$$\mathbb{E}_\psi(A^2)\mathbb{E}_\psi(B^2) \geq \frac{1}{4}(\mathbb{E}_\psi(C))^2 \tag{96}$$
with equality iif $(A - i\mu B)\psi = 0$ for some real μ.

Proof

The first one is obvious (just multiply); for the second

$$||(A - i\mu B)\psi||^2 \equiv \langle (A - i\mu B)\psi | (A - i\mu B)\psi\rangle = \langle \psi | (A - i\mu B)^* (A - i\mu B)\psi\rangle,$$

and use the first; now the LHS is non-negative so that the discriminant of the RHS as a quadratic in μ must be non-positive, which gives the third.

Corollary: Heisenberg's Uncertainty Principle

For normalized ψ,

$$\Delta_\psi(P)\Delta_\psi(X) \geq \frac{1}{2}\hbar \tag{97}$$

with equality iff

$$\psi = \text{const.} \times \exp\left(-\frac{\mu}{2\hbar}(x - \alpha)^2\right) \tag{98}$$

for some real μ and complex α.

Proof

Set $A = P - \mathbb{E}_\psi(P)I,\ B = X - \mathbb{E}_\psi(X)I$ so these are self-adjoint and

$$[A, B] = [P, X] = -i\hbar I$$

so take $C = -\hbar I$. Note that

$$\mathbb{E}_\psi(A^2) = (\Delta_\psi(P))^2, \quad \mathbb{E}_\psi(B^2) = (\Delta_\psi(X))^2, \quad \mathbb{E}_\psi(C) = -\hbar,$$

then the result follows from (96). For equality we need ψ, μ with

$$(A - i\mu B)\psi = 0.$$

Introduce $\beta = \mathbb{E}_\psi(P) - i\mu\mathbb{E}_\psi(X)$, then

$$(A - i\mu B)\psi = (-i\hbar\frac{d}{dx} - i\mu x - \beta)\psi = 0$$

which is a separable ODE leading to

$$\log\psi = -\frac{1}{2\hbar}(\mu x^2 - 2i\beta x + \text{const.})$$

which is (98) with $\alpha = i\beta/\mu$.

We shall call a state with wave-function like (98) a *minimum uncertainty state*. You saw an example in Part A, the ground state of the harmonic oscillator which had

$$\Psi_0(x) = \text{const.} \times \exp(-\frac{m\omega x^2}{2\hbar})$$

4.3 Simultaneous Measurability

Why is (97) called the Heisenberg Uncertainty principle? It implies that a reduction in the dispertion of P, $\Delta_\psi(P)$, must be accompanied by an increase in the dispersion of X and vice versa - there is a fundamental limit on reducing the common uncertainty. Thus P and X cannot be simultaneously be measured to arbitrarily high accuracy, and the origin of this is in the nonvanishing of the commutator. The same will hold for any pair of observables which don't commute.

What if $[A, B] = 0$? For a finite-dimensional $\mathcal{H}$ there will then be a basis of common eigenvectors, and on a common eigenvector A and B can be simultaneously measured exactly, i.e. the dispersions vanish. For an infinite-dimensional $\mathcal{H}$ one can do less. If A, B are self-adjoint introduce:

$\mathcal{H}_A$ = span of eigenvectors of A.

$\mathcal{H}_B$ = span of eigenvectors of B.

$\mathcal{H}_{AB}$ = span of common eigenvectors.

Then **Claim**: If $AB = BA$ then $\mathcal{H}_{AB} = \mathcal{H}_A \cap \mathcal{H}_B$

Then we can say the following:

Corollary

If $\mathcal{H}$ has an orthogonal basis of eigenvectors of A (so $\mathcal{H}_A = \mathcal{H}$) and $[A, B] = 0$ then $\mathcal{H}_{AB} = \mathcal{H}_B$ so that A, B can be simultaneously diagonalized on $\mathcal{H}_B$.

We shall see below, in section 5, some important application of this observation. There we shall have operators which commute with the Hamiltonian, so that energy eigenstates (i.e. eigenstates of the Hamiltonian) can simultaneously be eigenstates of these other operators. A result which can usefully be given here concerns operators which *do not* commute with the Hamiltonian. We have

Proposition

For any observable A write $\langle A \rangle = \langle \psi | A\psi \rangle$ then

$$\frac{d}{dt}\langle A \rangle = \langle \frac{\partial A}{\partial t} \rangle + \frac{i}{\hbar}\langle [H, A] \rangle \tag{99}$$

This follows rapidly from (66) by differentiating $\langle \psi | A\psi \rangle$. For an A with non explicit time-dependence the first term is zero, so that the rate of change of $\langle A \rangle$ is determined by the commutator $[H, A]$.

4.4 The Harmonic Oscillator Revisited

Now we come to the pay-off for this abstraction. Recall that the harmonic oscillator has Hamiltonian

$$H = \frac{1}{2m}P^2 + \frac{1}{2}m\omega^2 X^2 \tag{100}$$

Inspired by section 4.2 above, define

$$a_- = P - im\omega X \tag{101}$$

This operator is not self-adjoint and in fact its adjoint is

$$a_+ = (a_-)^* = P + im\omega X \tag{102}$$

These operators have the following important properties:

Lemma

1.
$$[a_-, a_+] = 2m\hbar\omega I$$
$$a_+a_- = 2m(H - \frac{1}{2}\hbar\omega I)$$
$$a_-a_+ = 2m(H + \frac{1}{2}\hbar\omega I)$$

2.
$$||a_-\psi||^2 = 2m(\mathbb{E}_\psi(H) - \frac{1}{2}\hbar\omega)||\psi||^2$$
$$||a_+\psi||^2 = 2m(\mathbb{E}_\psi(H) + \frac{1}{2}\hbar\omega)||\psi||^2$$

3.
$$[H, a_-] = -\hbar\omega a_-$$
$$[H, a_+] = \hbar\omega a_+$$

Proof: 1 is straightforward; for 2 write $\langle a_-\psi|a_-\psi\rangle = \langle\psi|a_+a_-\psi\rangle$ and use 1; for 3 use 1.

Proposition: The Harmonic Oscillator done by Algebra

Suppose H from (100) has a normalized eigenvector $\psi \;\epsilon\; \mathcal{H}$ then

1. there is a normalizable ground state ψ_0 with energy eigenvalue $E_0 = \frac{1}{2}\hbar\omega$;
2. there are normalizable excited states $\psi_n = (a_+)^n\psi_0$ for each positive integer n, with energy eigenvalue $E_n = (n + \frac{1}{2})\hbar\omega$;
3. if we take $P = -i\hbar\frac{d}{dx}$ and $X = x$ then $\psi_0 = \text{const.}\exp(-\frac{m\omega x^2}{2\hbar})$ and the $\psi_n, n = 0, 1, 2, ...$ are all the eigenstates.

Proof

Suppose the given normalized eigenvector satisfies

$$H\psi = E\psi$$

then

$$Ha_-\psi = (Ha_- - a_-H + a_-H)\psi = (a_-H + [H, a_-])\psi = (E - \hbar\omega)a_-\psi$$

by part 3 of the Lemma. That means that $a_-\psi$ is also an eigenvector of H, but with eigenvalue $(E - \hbar\omega)$. i.e. a_- "lowered" the eigenvalue by $\hbar\omega$. Also

$$||a_-\psi||^2 = 2m(\mathbb{E}_\psi(H) - \frac{1}{2}\hbar\omega)||\psi||^2 = 2m(E - \frac{1}{2}\hbar\omega)||\psi||^2$$

by part 2 of the lemma, so that this new eigenvector is normalizable.

We can then apply a_- again. We repeat this process obtaining normalizable eigenstates $(a_-)^k\psi$ with eigenvalue $E_k = E - k\hbar\omega$. However, this process must stop as

$$||(a_-)^k\psi||^2 \propto (E - (k - 1/2)\hbar\omega),$$

which cannot be negative. Thus the process will stop when we reach ψ_0 with $E = E_0 = \frac{1}{2}\hbar\omega$ and

$$||a_-\psi_0||^2 = 0, \quad \text{i.e. } a_-\psi_0 = 0$$

so that further lowering gives zero. (At this point we don't know the dimension of $ker\ a_-$, but we do know that any element of it has $E = E_0 = \frac{1}{2}\hbar\omega$). Now we start to 'raise': set $\psi_1 = a_+\psi_0$ then

$$H\psi_1 = (Ha_+ - a_+H + a_+H)\psi_0 = (\hbar\omega + \frac{1}{2}\hbar\omega)a_+\psi_0 = \hbar\omega(1 + 1/2)\psi_1$$

Inductively, raising will give normalizable states $\psi_n = a_+\psi_{n-1}$ with eigenvalues $E_n = (n + 1/2)\hbar\omega$. This process can be repeated indefinitely and we have a chain of eigenstates, labelled by the positive integers, arising from each candidate ψ_0. Every eigenstate must be in one of these chains, since the only elements of $ker\ a_-$ are lowest-energy states so that a_+ and a_- are one-to-one going up and down the chain.

As regards ψ_0 if we make the usual identification $P = -i\hbar\frac{d}{dx}$ and $X = x$, then we have

$$0 = a_-\psi_0 = (P - im\omega X)\psi_0 \ \Rightarrow\ (-i\hbar\frac{d}{dx} - im\omega x)\psi_0 = 0$$

from which

$$\psi_0 = \text{const.} \times \exp\left(-\frac{m\omega x^2}{2\hbar}\right)$$

This gives a unique normalized ground state ψ_0, i.e. in this case $dim(\ ker\ a_-) = 1$, and every other state is obtained by raising this one.

4.5 Uniqueness of the Commutation Relations

This is a technical aside: the algebraic approach of the previous subsection started just from the commutation relations

$$[P, X] = -i\hbar$$

We only used the identification $P = -i\hbar\frac{d}{dx}$ and $X = x$ to find $\psi_0(x)$ and hence $\psi_n(x)$. What freedom do we have in making this identification for P and X? If we had instead proceeded by assuming that $ker(a_-)$ was one-dimensional, spanned by Ω say, then the normalizable eigenstates would be $\psi_n^\Omega = (a_+)^n\Omega$. Now we would have the Hilbert space $\mathcal{H}^\Omega$ spanned by these and there would be an isomorphism of Hilbert spaces defined by relating the bases:

$$\psi_n^\Omega \to \psi_n(n)$$

Under this isomorphism the abstract P, X acting on $\mathcal{H}^\Omega$ would map to the operators $-i\hbar\frac{d}{dx}$, x acting on our previous Hilbert space of functions. Thus the identification $P = -i\hbar\frac{d}{dx}$ and $X = x$ is unique up to isomorphism.

5 Angular Momentum

(Hannabus chapter 8)

5.1 Angular Momentum in Quantum Mechanics

Motivated by the definition of angular momentum from classical mechanics:

$$\vec{L} = \vec{x} \wedge m\vec{v} = \vec{x} \wedge \vec{p} \tag{103}$$

we define in quantum mechanics the observables

$$\begin{aligned} L_1 &= X_2P_3 - X_3P_2 \\ L_2 &= X_3P_1 - X_1P_3 \\ L_3 &= X_1P_2 - X_2P_1 \end{aligned} \tag{104}$$

We call this $\vec{L}$ *orbital angular momentum.* More briefly

$$L_i = \epsilon_{ijk}X_jP_k \tag{105}$$

using the Einstein summation convention (repeated indices are automatically summed over their whole range) and in terms of the *alternating symbol* defined by

$$\begin{aligned}\epsilon_{ijk} &= 1 \text{ if } ijk \text{ is an even permutation of } 123\\ &= -1 \text{ if } ijk \text{ is an odd permutation of } 123\\ &= 0 \text{ otherwise}\end{aligned}$$

(so in particular $(\vec{a} \wedge \vec{b})_i = \epsilon_{ijk} a_j b_k$)

Properties

1. $[L_i, P_j] = i\hbar\epsilon_{ijk} P_k$;
2. $[L_i, X_j] = i\hbar\epsilon_{ijk} X_k$;
3. $[L_i, L_j] = i\hbar\epsilon_{ijk} L_k$

Proof

We'll just do the first of these. One can do them in an unsophisticated way, calculating e.g. $[L_1, P_2]$ and appealing to symmetry or in a more sophisticated way as follows:

$$[L_i, P_j] = [\epsilon_{imn} X_m P_n, P_j] = \epsilon_{imn}[X_m, P_j] P_n = i\hbar\epsilon_{imn}\delta_{mj} P_n = i\hbar\epsilon_{ijn} P_n$$

as required. This needs some familiarity with the Einstein summation convention, a skill worth adquiring.

We shall call any set $\vec{J} = (J_1, J_2, J_3)$ of observables which satisfy

$$[J_i, J_j] = i\hbar\epsilon_{ijk} J_k \tag{106}$$

an angular momentum operator, and we will study them in the abstract.

More Properties

If $[J_i, A_j] = i\hbar\epsilon_{ijk} A_k$ and $[J_i, B_j] = i\hbar\epsilon_{ijk} B_k$ then

$$[J_i, \vec{A} \cdot \vec{B}] = 0 \tag{107}$$

Proof

$$[J_i, \vec{A} \cdot \vec{B}] = [J_i, A_j B_j] = [J_i, A_j] B_j + A_j [J_i, B_j] = i\hbar\epsilon_{ijk}(A_k B_j + A_j B_k) = 0$$

where the last step follows because ϵ_{ijk} changes sign under switching a pair of indices and $(A_k B_j + A_j B_k)$ does not.

We deduce in particular that $\vec{P} \cdot \vec{P}$, $\vec{P} \cdot \vec{X}$, $\vec{X} \cdot \vec{X}$, $\vec{L} \cdot \vec{L}$ and similar expressions all commute with $\vec{L}$.

5.2 Ladder Operators

We have just seen that

$$[J^2, J_i] = 0 \tag{108}$$

writing J^2 for $\vec{J} \cdot \vec{J}$. We define the *ladder operators*

$$J_{\pm} = J_1 \pm i J_2 \tag{109}$$

and note that $(J_+)^* = J_-$ assuming the J_i are self-adjoint.

Properties

1. $[J^2, J_{\pm}] = 0,$
2. $J_+J_- = J^2 - J_3^2 + \hbar J_3,$
   $J_-J_+ = J^2 - J_3^2 - \hbar J_3,$
   $[J_+, J_-] = 2\hbar J_3,$
3. $[J_3, J_{\pm}] = \pm\hbar J_{\pm}$

1 is done and 2 and 3 follow rapidly from the definitions.

Now, much as with the algebraic treatment of the Harmonic oscillator, we will find eigenvectors of J^2 and J_3. They commute so we will be looking for simultaneous eigenvectors. Assume these exist in some $\mathcal{H}$, so there is a ψ with

$$J^2\psi = \lambda\hbar^2\psi, \quad J_3\psi = m\hbar\psi,$$

for real λ, m. Consider the new states

$$\psi_+ = J_+\psi, \quad \psi_- = J_-\psi,$$

then

$$J^2\psi_- = J^2J_-\psi = (J^2J_- - J_-J^2 + J_-J^2)\psi = \lambda\hbar^2\psi_- \tag{110}$$

i.e. since $[J^2, J_-] = 0$, ψ_- is an eigenvector of J^2 with the same eigenvalue. This is true for ψ_+ too.

Next

$$J_3\psi_- = J_3J_-\psi = (J_3J_- - J_-J_3 + J_-J_3)\psi = (m-1)\hbar\psi_- \tag{111}$$

so the J_3-eigenvalue is lowered by one for ψ_- (and raised by one for ψ_+).

Next calculate

$$||\psi_-||^2 = \langle J_-\psi|J_-\psi\rangle = \langle\psi|J_+J_-\psi\rangle = (\lambda - m^2 + m)\hbar^2||\psi||^2 \tag{112}$$

which must be non-negative, as

$$||\psi_+||^2 = \langle J_+\psi|J_+\psi\rangle = \langle\psi|J_-J_+\psi\rangle = (\lambda - m^2 - m)\hbar^2||\psi||^2 \tag{113}$$

which also must be non negative.

Putting these last two results together, we have

$$\begin{aligned} &\lambda \geq m(m-1) \text{ and equality iff } J_-\psi = 0 \\ &\lambda \geq m(m+1) \text{ and equality iff } J_+\psi = 0 \end{aligned} \tag{114}$$

This is enough to obtain the results below.

5.3 Representation of the Angular Momentum

Here 'representations' means as operators acting on some explicit vector space, so we suppose we have a finite-dimensional, complex vector space $\mathcal{H}$ of eigenvectors of J^2:

$$J^2\psi = \lambda\hbar^2\psi \text{ for all } \psi \in \mathcal{H}.$$

J_3 will have an eigenvector in $\mathcal{H}$ say ψ with

$$J_3\psi = m\hbar\psi$$

We can obtain more eigenvectors of J_3 by raising and lowering with $J_\pm$, when m will go up and down in integer steps. This can't be repeated indefinitely, since (114) implies

$$\lambda + \frac{1}{4} \geq m^2 \pm m + \frac{1}{4} = (m \pm \frac{1}{2})^2$$

so that $|m|$ can't be arbitrarily large. Thus there will be a maximum m, call it j with corresponding ψ_j, and further raising must give zero:

$$J_+\psi_j = 0$$

By (113), if the LHS vanishes, we find

$$\lambda = j(j+1)$$

We lower ψ_j in integer steps and this will stop at a value of m at which further lowering gives zero. By (112) this will happen when

$$0 = \lambda - m^2 + m = j(j+1) - m^2 + m = (j+m)(j-m+1)$$

i.e. at $m = -j$ since $m = j + 1$ is not allowed (m must be less than its maximum value which was j)

Thus we lower in integer steps from $m = j$ to $m = -j$, which requires $2j$ to be a non-negative integer.

Summarizing:

- the eigenvalues of J^2 are $j(j+1)\hbar^2$ where $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$;
- for each fixed j, the eigenvalues of J_3 are $m\hbar$ with $m = -j, -j+1, \ldots, j-1, j$;
- assuming no degeneracies, the eigenspace on which $J^2 = j(j+1)\hbar$ has dimension $2j+1$.
- we can choose an orthonormal basis $\{\psi_m\}$ whith

$$\begin{aligned} J_3\psi_m &= m\hbar\psi_m \\ J_\pm\psi_m &= c_m^\pm\psi_{m\pm1} \\ c_m^\pm &= \left[(j \mp m)(j \pm m + 1)\right]^{1/2} \end{aligned}$$

We have proved all of this except the last part which comes from (113) and (112).

Important example: $j = \frac{1}{2}$ (also called the spin representation)

Now $2j + 1 = 2$ so $\mathcal{H}$ is 2-dimensional. The allowed values of m are $m = \pm\frac{1}{2}$ so write $\psi_\pm$ for $\psi_{\pm1/2}$, then

$$J_3\psi_\pm = \pm\frac{1}{2}\hbar\psi_\pm \tag{115}$$

and

$$J_+\psi_+ = J_-\psi_- = 0, \qquad J_+\psi_- = \hbar\psi_+, \qquad J_-\psi_+ = \hbar\psi_- \tag{116}$$

We can introduce a matrix formalism by taking as basis

$$\psi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \psi_+ = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Then (115) and (116) fix the form of the operators to be

$$J_3 = \frac{1}{2}\hbar\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad J_+ = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad J_- = \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

so also

$$J_1 = \frac{1}{2}\hbar\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad J_2 = \frac{1}{2}\hbar\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

and one may check all the commutation relations. These formulae bring to our attention the *Pauli spin matrices*:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{117}$$

which are Hermitian, trace zero and define the J_i for $j = 1/2$ via

$$J_i = \frac{1}{2}\hbar\sigma_i \tag{118}$$

from which many algebraic relations follow.

5.4 Orbital Angular Momentum and Spherical Harmonics

Let us focus in the particular case of orbital angular momentum $\vec{J} = \vec{L} = \vec{X} \wedge \vec{P}$. We seek a space of functions for $\vec{L}$ to act on. As we will shortly see, we need j to be an integer. To justify this claim recall

$$L_3 = X_1P_2 - X_2P_1 = -i\hbar(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial_x}) = -i\hbar\frac{\partial}{\partial\phi} \tag{119}$$

in terms of spherical polar coordinates. Now the eigenvalue equation becomes

$$L_3\psi = -i\hbar\frac{\partial\psi}{\partial\phi} = m\hbar\psi \tag{120}$$

so that

$$\psi = F(r,\theta)e^{im\phi} \tag{121}$$

Now the important point: if this is genuinely a function, i.e. single-valued, then we need m to be an integer, so j is also an integer. (We also see that $j = 1/2$ must be associated with something, not a function, which changes sign under rotation by 2π.) To pursue this line further, lets calculate the other operators in spherical polar coordinates

$$\begin{aligned} L_+ &= \hbar e^{i\phi}\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right) \\ L_- &= -\hbar e^{-i\phi}\left(\frac{\partial}{\partial\theta} - i\cot\theta\frac{\partial}{\partial\phi}\right) \\ L^2 &= -\hbar^2\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right) \end{aligned} \tag{122}$$

We will find the eigenfunctions explicitly, so write $\psi_{\ell m}$ for the one with

$$L^2\psi_{\ell m} = \ell(\ell+1)\hbar^2\psi_{\ell m}, \quad L_3\psi_{\ell m} = m\hbar\psi_{\ell m} \tag{123}$$

then the highest m-value is ℓ and $\psi_{\ell\ell}$ satisfies

$$L_3\psi_{\ell\ell} = -i\hbar\frac{\partial\psi_{\ell\ell}}{\partial\phi} = \ell\hbar\psi_{\ell\ell}, \qquad L_+\psi_{\ell\ell} = 0 \tag{124}$$

The first is solved by $\psi_{\ell\ell} = F(\theta)e^{i\ell\phi}$, where we are only looking at the θ, ϕ dependence (the r-dependence will be reinstated in the next section), and then the second is

$$L_+\psi_{\ell\ell} = \hbar e^{i\phi}\left(\frac{dF}{d\theta} - \ell(\cot\theta)F\right)e^{i\ell\phi} = 0$$

whence $F = \text{const.} \times (\sin\theta)^\ell$, so

$$\psi_{\ell\ell} = \text{const.} \times (\sin\theta)^\ell e^{i\ell\phi}. \tag{125}$$

The constant can be fixed by normalization. Once we have this we obtain the other $\psi_{\ell m}$ by lowering with L_-. We shall write $Y_{\ell m}(\theta, \phi)$ for the corresponding functions. Since they arose as eigenfunctions we know at once that

$$\langle\psi_{\ell_1 m_1}|\psi_{\ell_2 m_2}\rangle = \int_0^{2\pi}\int_0^{\pi}\overline{Y_{\ell_1 m_1}}Y_{\ell_2 m_2}\sin\theta d\theta d\phi = 0, \quad \text{if } (\ell_1, m_1) \neq (\ell_2, m_2). \tag{126}$$

5.5 The Hydrogen Atom Again

With a better understanding of angular momentum we can do a more thorough treatment of the hydrogen atom. The time-independent Schrodinger equation for a particle of mass m and charge $-e$ (the electron) in the electric field of a fixed point of charge Ze (the nucleus, so $Z = 1$ for the hydrogen atom) is

$$-\frac{\hbar^2}{2m}\nabla^2\Psi - \frac{Ze^2}{4\pi\epsilon_0 r}\Psi = E\Psi \tag{127}$$

and the Laplacian in spherical coordinates is

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right) \tag{128}$$

From (122), we recognize that the angular derivatives are proportional to L^2, so that we can separate $\Psi = F(r)Y_{\ell m}(\theta, \phi)$. We then obtain

$$\nabla^2\Psi = \left(\frac{1}{r}(rF)'' - \frac{\ell(\ell+1)}{r^2}F\right)Y_{\ell m} \tag{129}$$

writing prime for d/dr.

We simplify the Schrodinger equation by renaming quantities:

$$E = -\frac{\hbar^2\kappa^2}{2m}, \qquad \beta = \frac{2m}{\hbar^2}\frac{Ze^2}{4\pi\epsilon_0}, \tag{130}$$

Further, we write $F(r)$ in terms of a new function:

$$F(r) = f(r)e^{-\kappa r} \tag{131}$$

We find that f satisfies the following equation

$$f'' + \left(\frac{2}{r} - 2\kappa\right) f' - \left(\frac{2\kappa - \beta}{r} + \frac{\ell(\ell+1)}{r^2}\right) f = 0 \tag{132}$$

We seek a series solution

$$f = \sum_{n=0}^{\infty} a_n r^{n+c}, \qquad \text{with } a_0 \neq 0, \tag{133}$$

and substitute into (132) to obtain

$$\sum \left(a_n(n+c)(n+c-1)r^{n+c-2} + \left(\frac{2}{r} - 2\kappa\right) a_n(n+c)r^{n+c-1} - \left(\frac{2\kappa-\beta}{r} + \frac{\ell(\ell+1)}{r^2}\right) a_n r^{n+c} \right) = 0 \tag{134}$$

the lowest power is r^{c-2} and the coefficient is

$$a_0(c(c-1) + 2c - \ell(\ell+1))$$

This must vanish and $a_0 \neq 0$ so we obtain the *indicial equation*:

$$(c(c-1) + 2c - \ell(\ell+1)) = (c-\ell)(c+\ell+1) = 0 \tag{135}$$

The root $c = -(\ell+1)$ is not allowed as this would give Ψ unbounded at $r = 0$, thus $c = \ell$. For the general power r^{n+c-1} we obtain the *recurrence relation*:

$$a_{n+1}((n+c+1)(n+c) + 2(n+c+1) - \ell(\ell+1)) = a_n(2\kappa(n+c) + 2\kappa - \beta), \tag{136}$$

which determines all terms in the series from a_0. If the series continues to infinity, then for large n

$$\frac{a_{n+1}}{a_n} \approx \frac{2\kappa}{n}$$

and so

$$f \approx e^{2\kappa r}$$

Tracking back through (131), we see that this will give a non-normalizable Ψ. So normalizability requires the series to terminate and from the recurrence relation there must be a non-negative n' for which the RHS vanishes, i.e.

$$2\kappa(n' + \ell + 1) = \beta \tag{137}$$

where we have also substituted for c.

Set $n = n' + \ell + 1$, which is a positive integer, then the series for f takes the form

$$f = f_n = r^\ell \sum_{k=0}^{n'} a_k r^k, \tag{138}$$

which is a polynomial of degree $n-1$, with corresponding

$$\kappa = \kappa_n = \frac{\beta}{2n}. \tag{139}$$

The wave function is

$$\Psi_{n\ell m} = \text{const.} \times f_n(r) e^{-\kappa_n r} Y_{\ell m}(\theta, \phi) \tag{140}$$

and the energy eigenvalue is

$$E = E_n = -\frac{\hbar^2 \kappa_n^2}{2m} = -\frac{\hbar^2 \beta^2}{8m} \frac{1}{n^2} \tag{141}$$

The integers n, ℓ, m, which label the state, are known respectively as the *principal quantum number*, the *azimuthal quantum number* and the *magnetic quantum number*. The energy eigenvalue notices only n, not ℓ or m. Any spherically symmetric potential will ignore m, which only comes into E when spherical symmetry is given up, for example by putting the atom in a magnetic field (whence the name), but spherical potentials other than the pure Coulomb (i.e. $V \approx \frac{1}{r}$) will have ℓ appearing in E.

We can count the degeneracy of an energy eigenvalue as follows: given $n = n' + \ell + 1$ we have $0 \leq \ell \leq n-1$ and for each ℓ there are $2\ell + 1$ choices for m, so that the degeneracy is

$$\sum_{\ell=0}^{n-1} (2\ell + 1) = n^2$$

6 A Charged Particle in a Magnetic Field

(this section is not part of the course and so not examinable, but it makes a link back to the material on electromagnetism, and introduces some interesting new ideas.)

To connect the second half of the course with the first, we briefly consider the quantum version of a classical charged particle of mass m and charge e moving in a electromagnetic field. In the case of an electric field, with electric potential say Φ, we know how to do this: simply consider the Hamiltonian

$$H = \frac{1}{2m} |\vec{P}|^2 + V(\vec{X})$$

with $V(\vec{X}) = e\Phi(\vec{X})$. The hydrogen atom is just like this.

In the case of a magnetic field, recall that the force does not work on the particle, so does not contribute a potential energy term: something new is needed. To motivate what this is, and because it gives some insight into how classical mechanics can be recovered from quantum mechanics, look at (99):

$$\frac{d}{dt}\langle A\rangle = \langle\frac{\partial A}{\partial t}\rangle + \frac{i}{\hbar}\langle[H,A]\rangle \tag{142}$$

With the Hamiltonian above, this will give

$$\frac{d}{dt}\langle\vec{X}\rangle = \frac{i}{\hbar}\langle[H,\vec{X}]\rangle = \frac{i}{\hbar}\langle[\frac{1}{2m}|\vec{P}|^2,\vec{X}]\rangle = \frac{1}{m}\vec{P},$$

and

$$\frac{d}{dt}\langle\vec{P}\rangle = \frac{i}{\hbar}\langle[H,\vec{P}]\rangle = \frac{i}{\hbar}[V(\vec{X}),\vec{P}] = -\langle\nabla V\rangle$$

Combining these two we obtain

$$m\frac{d^2}{dt^2}\langle\vec{X}\rangle = -\langle\nabla V\rangle \tag{143}$$

which closely parallels the classical equations of motion for a particle in a potential, which read

$$m\frac{d^2}{dt^2}\vec{x} = -\nabla V \tag{144}$$

In other words, the expectation values satisfy the classical equations of motion. Now we consider the Hamiltonian

$$H = \frac{1}{2m}\left((P_1+\frac{1}{2}eBY)^2 + (P_2-\frac{1}{2}eBX)^2 + P_3^2\right) \tag{145}$$

which is similar to one introduced on one of the questions sheets, where we found the energy levels without the term P_3^2. What classical problem does this corresponds to? We find

$$\frac{d}{dt}\langle X\rangle = \frac{i}{\hbar}\langle[H,X]\rangle = \frac{1}{m}\langle P_1+\frac{1}{2}eBY\rangle \equiv \langle V_1\rangle \tag{146}$$

where the right-hand-side serves to define the "velocity" V_1. Similarly

$$\frac{d}{dt}\langle Y\rangle = \frac{i}{\hbar}\langle[H,Y]\rangle = \frac{1}{m}\langle P_2-\frac{1}{2}eBX\rangle \equiv \langle V_2\rangle \tag{147}$$

while

$$\frac{d}{dt}\langle Z\rangle = \frac{i}{\hbar}\langle[H,Z]\rangle = \frac{1}{m}\langle P_3\rangle \equiv \langle V_3\rangle \tag{148}$$

Then

$$
\begin{aligned}
\frac{d}{dt}\langle P_1\rangle &= \frac{i}{\hbar}\langle [H,P_1]\rangle = \frac{1}{2}eB\langle V_2\rangle \\
\frac{d}{dt}\langle P_2\rangle &= \frac{i}{\hbar}\langle [H,P_2]\rangle = -\frac{1}{2}eB\langle V_1\rangle \\
\frac{d}{dt}\langle P_3\rangle &= 0
\end{aligned}
$$

so that

$$
\begin{aligned}
m\frac{d^2}{dt^2}\langle X\rangle &= eB\langle V_2\rangle \\
m\frac{d^2}{dt^2}\langle Y\rangle &= -eB\langle V_1\rangle \\
m\frac{d^2}{dt^2}\langle Z\rangle &= 0
\end{aligned}
\tag{149}
$$

This system corresponds to the classical equation

$$
m\frac{d^2}{dt^2}\vec{x} = e\vec{v}\wedge\vec{B} \tag{150}
$$

with $\vec{B} = B\vec{k}$, which is the equation of motion of a classical particle of mass m and charge e moving in the magnetic field $\vec{B}$, which is uniform and in the z-direction

$$
\vec{B} = (0,0,B)
$$

To see how to get (145) note that this magnetic field has vector potential

$$
\vec{A} = \frac{1}{2}B(-y,x,0) \tag{151}
$$

so that (145) may be obtained by starting with the free particle Hamiltonian $H = \frac{1}{2m}|\vec{P}|^2$ and making the replacement

$$
\vec{P} \to \vec{P} - e\vec{A} \tag{152}
$$

It is a straightforward calculation to see that this is the correct prescription more generally so that one could, for example, go on to calculate the energy levels of a hydrogen atom in a uniform magnetic field. We won't do that, but to end we will note a new interpretation to gauge transformations that this gives. Recall a gauge transformation was a change in $\vec{A}$:

$$
\vec{A} \to \vec{\tilde{A}} = \vec{A} + \nabla\zeta \tag{153}
$$

which leaves the magnetic field unchanged. This should therefore lead to no physical change, but it will change the term $\frac{1}{2m}|\vec{P} - e\vec{A}|^2$ in H. The resolution of this puzzle is to make a simultaneous change in the wave function:

$$\psi \to \tilde{\psi} = \psi \exp(ie\zeta/\hbar) \tag{154}$$

This change doesn't affect the probability density ρ or probability current $\vec{j}$, which are the physical observable quantities, and also

$$|\vec{P} - e\tilde{\vec{A}}|^2\tilde{\psi} = \exp(ie\zeta/\hbar)|\vec{P} - e\vec{A}|^2\psi \tag{155}$$

so that the term $\exp(ie\zeta/\hbar)$ factors out of the Schrodinger equation, and solutions will transform to solutions.